PHILIP MORRIS U.S.A.

INTEROFFICE CORRESPONDENCE

Richmond, Virginia

To: **Distribution** **Date:** July 21, 2000

From: **Cindy Boswell, Ron Parker**

Subject: **Evaluation of Measurement Uncertainty for Filtrona FTC Smoking Results**

Objective. To establish the uncertainty values for smoking parameters (TPM, Puff Count, CO, Nicotine, Water, and Tar) for results generated off of Product Testing Laboratory (PTL) Filtrona smoking machines and HP5890 GCs, under FTC conditions.

Summary. The following results were obtained as estimates of expanded uncertainty (with a coverage factor of 3) for FTC Filtrona smoking in PTL:

Parameter	Product's Nominal Tar Level	Number of Ports	Expanded Uncertainty
Puff Count	1.0-8.0	4	0.4
	1.0-8.0	8	0.3
Carbon Monoxide (CO)	16.0	4	1.4
	16.0	8	1.0
	8.0	8	0.9
	1.0	8	0.3
Total Particulate Matter (TPM)	16.0	4	2.3
	16.0	8	1.6
	8.0	8	0.8
	1.0	8	0.5
Nicotine (NIC)	16.0	4	0.09
	16.0	8	0.06
	8.0	8	0.06
	1.0	8	0.02
Water	16.0	4	1.12
	16.0	8	0.79
	8.0	8	0.21
	1.0	8	0.13
TAR	16.0	4	1.6
	16.0	8	1.2
	8.0	8	0.7
	1.0	8	0.5

Background. The "Guide to the Expression of Uncertainty in Measurement"[1] states that all factors affecting measurement variance be included in an evaluation of uncertainty, and each of those causes be identified as either Type A or Type B. Type A uncertainties are those which can be evaluated by statistical analysis of observations. This can be interpreted to mean that they are in statistical control and will be accounted for in a well-designed study of the measurement process. Type B uncertainties are those that are evaluated by means other than statistical analysis of observations. Some of the Type B uncertainties may be provided by the supplier of the equipment or input. Type B uncertainties must then be propagated into the expanded (overall) uncertainty, following the law of propagation of errors. Experience has shown that propagation often leads to unreasonable and impractical estimates of expanded uncertainty. A solution is to attempt to prove that all influence factors can be addressed as Type A uncertainties, alleviating the need for propagating uncertainties. That is, those factors that are in statistical control are predictable and consistent with normal variation and it can be stated that their influence on the results is incorporated in the overall estimate of variation. It is likely that, despite best efforts, a few Type B uncertainties will remain and they can be propagated into the final expanded uncertainty.

Method. The smoking process is among the most complex. There are several elements that must be included. Each smoking machine must be tested for several runs over an extended period. Brands of several different nominal tar levels are tested on each machine. Several operating parameters and environmental conditions must be observed and recorded during the length of the study. All TPM pads from a single smoking machine are analyzed on the same GC to attempt to minimize the effect of the GC portion of the process. A separate analysis of the GC operations is then performed to isolate its contribution to the overall uncertainty. The outcome of this type of study may necessitate a second trial to incorporate new learnings.

Experimental Design.

Samples – Utilize several cigarette models to cover the desired tar range. For this experiment, two production brands were utilized (Merit KS with nominal tar level 8.0 mg/cigt. and Next with nominal tar level 1.0 mg/cigt.) and the Industry Monitor #16 (IM#16) with nominal tar level 16.0 mg/cigt. The Merit and Next brands were specially requested from the factory to obtain batches of sample produced in a short run to minimize sample variation. The IM#16 product was retrieved from cold storage.

Machines – Utilize all involved smoking machines.

Test Volume – On each machine, smoke one run, consisting of 4 (IM#16) or 8 (regularly manufactured brands) ports of each brand, at least 3 times a week for at least 3 consecutive weeks. The IM#16 samples were always placed in their normal ports as designated by PTL's procedures. The position of the other two samples was alternated throughout the ports.

Conditions – Maintain and monitor conditions, recording temperature, humidity, and barometric pressure, during both equilibration and smoking.

Operating Parameters – Record airflow and puff volume and only adjust to correct an out of range condition (which will be recorded).

Operators – Utilize multiple trained personnel to operate the machines, rotating among the machines for each run. Maintain detailed logs for each run, recording any equipment failures and corrective action, operator errors, and unusual observations.

Plan – The smoking plan used for the first trial of this study was as follows:

---

[1] *Guide to the Expression of Uncertainty in Measurement.* International Oganization for Standardization. Switzerland. 1995.

		Filtrona Machine			
Run #	Week/Day	1	2	3	4
1	1/1	Oper A	Oper B	Oper A	Oper B
2	1/2	Oper B	Oper A	Oper A	Oper B
3	1/3	Oper A	Oper B	Oper B	Oper A
4	2/1	Oper B	Oper A	Oper B	Oper A
5	2/2	Oper B	Oper B	Oper A	Oper A
6	2/3	Oper A	Oper A	Oper B	Oper B
7	3/1	Oper A	Oper B	Oper A	Oper B
8	3/2	Oper B	Oper A	Oper A	Oper B
9	3/3	Oper A	Oper B	Oper B	Oper A

For the second trial, a similar plan was followed with the exception that the ninth run was smoked under different conditions for a separate experiment. The data from that run were not used in the analysis of these results.

Data Analysis. Once the study is complete, all data must be analyzed to determine:

- whether the various operating parameters and environmental conditions were in states of statistical control (control charts),
- whether there are significant differences among the machines (analysis of variance), and
- whether there are parameters that were not taken under consideration that need to be addressed.

Once the data are complete and ready for final calculations, the Type A uncertainty can be estimated from the variation in the results (specific to given nominal tar levels, if differences occur). Type B uncertainties, if any, can be propagated in with the Type A uncertainty to provide an expanded uncertainty.

Results. By creating a dependency diagram for the Filtrona SM350/400 FTC smoking process (Attachment 1), possible causes of variation were identified. Each possible cause was evaluated as to whether it introduces uncertainty into the result or invalidates results (Attachment 2). After the first three-week trial, it was determined that an additional trial was necessary to further investigate several items and to examine a possible long-term time effect (trial-to-trial bias or shift). After each of the two trials, all variable causes were control charted (Attachment 3), exceptions to stability (out-of-control incidences) were noted, investigated, and acted upon if appropriate (Attachment 4). After excluding invalid results based on certain out-of-control incidences, all variable causes were determined to be in control. Therefore, all of these potential uncertainty factors could be considered Type A and already incorporated in the overall estimate of variation. In other words, none of these causes need be addressed as Type B uncertainties and propagated in with the Type A estimate.

All product measurements (TPM, Puff Count, CO, Nicotine, Water, and Tar) were control charted two ways. First, standard deviation charts were produced to assess the between-port or within-run stability for each machine. These charts enabled identification of out-of-control data points that were often suspected to be sample mix-ups or GC-related errors. These were evidenced by a standard deviation of a run being abnormally high. Often, examination of the

individual data values constituting the high standard deviation would highlight one value that was either much higher or much lower than the others. Since each value is a composite of five smoked cigarettes, examination of the individual puff counts from the smoking machine output would show one of the five cigarettes to be either much higher or much lower than the others, indicating a single cigarette mix-up. Out of approximately 1400 data values, 20 were determined to be invalid and were excluded from further analysis. All charts in Attachment 5 were produced after these invalid data points were excluded. Attachment 6 details the average standard deviation for each product measurement for each brand within each machine over the course of the two trials.

The second method for control charting the product measurements was with individual and moving range charts. For this application, the average of each set of 4 or 8 ports for a sample within a run on a machine was considered a single result (Attachment 7). These "run-to-run" charts enabled an evaluation of the stability of the overall process within a machine. In general, these charts suggested that each machine was stable within each trial. Several charts showed out-of-control occurrences resulting from a shift in values from trial one to trial two. For the two production brands, these shifts could be attributed in part to the fact that a different batch of product was obtained for each of the two trials. The shifts observed in the IM#16 results suggest a time effect, or trial-to-trial bias. A few other out-of-control occurrences were noted and their causes were hypothesized (Attachment 8).

Attachment 9 details the summary statistics for all product measurements from the two trials. When comparing results from different machines, there were some incidences of statistically significant differences between means or variances, but they were not practically different enough to impact the uncertainty determinations. Since all causes of variation were deemed Type A, and all product measurements were in control after data validation (with the exception of a trial-to-trial bias that will be addressed later), we could now begin to estimate uncertainty based on measurement variation. Attachment 10 illustrates, for each product measurement parameter, a graph of the standard deviation against the average, along with a regression equation and correlation coefficient. This was an attempt to determine if the variation changes significantly as the average level changes significantly (presumably between the three different product levels). With the exception of Puff Count (R^2~0.11), all parameters showed a strong correlation between the average and the standard deviation (R^2>0.83). This suggests that Puff Count uncertainty need not be further qualified by Puff Count level (average). In addition, the Nicotine standard deviations for Merit and IM#16 were not statistically different. Therefore, those product levels were grouped for a Nicotine uncertainty estimate.

To estimate the Type A uncertainty, the standard deviations from each machine within each trial were used (reference Attachment 9). These standard deviations were converted to variances and then to the natural logarithms to obtain a composite average and standard deviation for each level so that an upper 95% confidence limits on the variance could be computed (Attachment 11). These upper 95% confidence limits were then converted back from the natural logarithm and back to standard deviations (from variances). These final values are the estimates of Type A uncertainty.

Trial-to-trial differences in parameter levels (averages) for the two production samples were attributed to the two separate batch productions. However, there were also differences noted in the IM#16 results, where all the product from both trials were produced at one time (Attachment 12). With the exception of Puff Count, all four machines showed increases in averages from the

first trial to the second trial. Although minor in magnitude, these increases indicate a statistically significant difference from trial one to trial two. This trial-to-trial effect is evidence of some sort of Type B uncertainty that will need to be propagated into the total, although the cause is unknown. We hypothesized that this Type B uncertainty should be smaller in magnitude than the Type A uncertainty because the degree of the observed mean shift was small. In addition, we presented a comprehensive inventory of causes of uncertainty and there is no historical evidence for additional causes whose effect would be much greater than controlled process variation. We also hypothesized that since the Type A uncertainty varies with product level, this Type B uncertainty should also vary with product level and therefore should be proportional to the Type A uncertainty. For lack of a scientific method, it was decided to compute Type B uncertainty as two-thirds of Type A uncertainty. Attachment 13 shows the resulting expanded uncertainties when Type A and Type B uncertainties are combined by taking the square root of the sum of the squares (propagation of errors) and a standard coverage factor of three is factored in. The discrimination is twice the expanded uncertainty, or the width of the uncertainty interval, and denotes the ability to detect differences between 4-port (IM#16) or 8-port (samples) averages. For example, when assessing Tar results of a product in the neighborhood of 8mg/cigt (like the Merit sample), an 8-port average can be expressed as average±0.7mg (uncertainty). When comparing one 8-port average to another, the two averages would have to be at least 1.3 mg/cigt apart for the intervals to not overlap and to claim that the averages are different. Attachment 14 converts these uncertainties to single-port references.

The two experimental trials and subsequent analyses were performed while minimizing the GC-related factors by utilizing one dispensing unit and one GC (with one backup) throughout the study. Dependency diagrams were created for several of the sub-processes (Attachment 15) and possible causes of variation were identified (Attachment 16). Our studies used a single Gilson dispensing unit. Actual dispensing variation of this dispenser is already accounted for in our final results, as we assume that the dispensing unit was operated in a correct and consistent fashion. However, the manufacturer's certificate for this devise reports the devices are accurate to ±2%. We take this to mean that the multiple dispensers will on average dispense quantities within a ±2% range. An underdispensing unit will make nicotine and other substance-in-solution calculations high, and an overdispensing unit will make our final results seem low. Since this variation is multiplicative, and we desire to report a conservative value, we have adjusted our "single dispenser" estimate of uncertainty by a multiplier of 1.00/0.98, or 1.0204. This adjustment is applied to all three levels of uncertainty estimates.

Three additional potential sources of variation were found to be significant. Lot-to-lot differences in the extract and stock solutions were evidenced by mean shifts in calibration slopes. These calibration slopes were plotted over a five-month time period that started prior to and ends after the experimental period (Attachment 17). In particular, some of the plots of the water calibration slopes clearly highlight different batches of solution. The mean shift was estimated to be approximately 3% for nicotine and 4% for water. The plots of calibration slopes also revealed some machine-to-machine differences in standard deviations (precision) when examining a period of stability for all machines (approximately the last two months studied). Therefore, the maximum standard deviation observed during this period on all ten machines (0.0033mg for nicotine and 0.0277 mg for water) was taken to be the estimate of Type B machine-to-machine precision uncertainty. Examination of check standard results revealed machine-to-machine differences in averages (accuracy). Therefore, the maximum difference between any two machines (0.0038mg for nicotine and 0.0526mg for water) was taken to be the estimate of Type B machine-to-machine accuracy uncertainty.

The composite uncertainty values for these three Type B factors equate to 0.03mg for nicotine and 0.15mg for water. The previous Type B uncertainty estimated from unknown sources in the smoking process (0.03mg for nicotine and 0.41mg for water) is equal to or greater than that calculated from difference in chemistry-related issues. Therefore, analysis of these chemistry-related Type B sources has validated the original Type B uncertainty and the only known additional source of Type B uncertainty is from dispenser variation.

Finally, using the estimate of dispenser variation for additional Type B uncertainty, the expanded uncertainty is recalculated on Attachment 18. The single port uncertainty values are on Attachment 19.

Attachments

Distribution:
Cindy Blair
Liz Chambers
Mary Ellen Counts
Bev Covitz
Carolyn Hendrick
Robin Holleman
Susan Laffoon
Judith Lightner
Wade Mokarry
Martha Payne
Jeff Sampson
Ed Sterling
Brenda Strang
Barbara Ward
Nita Winston